Rural Carrier Cost System – Digital DPS
Statistical Documentation

A. Overview

The Rural Carrier Cost System (RCCS) is a continuous, ongoing cross-sectional statistical study, or probability sample of rural carrier route-days. For each selected route-day, a sample of mail is selected, and for each selected mailpiece, the class, compensation category, shape, and other characteristics are recorded directly into a portable microcomputer using the Computerized On-Site Data Entry Systems (CODES) software. **RCCS Digital DPS** is proposed to be part of RCCS, with the major difference being a probability sample of rural carrier ZIP-days, and a substantial increase in sampled DPS mailpieces.

The RCCS gathers data for distributing major portions of carriers' salaries, benefits and related costs to the categories of mail for postal rate-making and related postal management purposes. Accrued carrier costs, available from payroll data, are aggregate amounts and are not generally associated with any particular class of mail or service. Therefore, special methods are needed to determine the costs associated with the mail categories.

The RCCS proposes to use data from Origin-Destination Information System – Revenue, Pieces, and Weight (ODIS-RPW) digital samples destined for delivery by rural carriers to enhance the estimation of delivered DPS volumes and replace a large portion of manual sampling of DPS letter trays by RCCS data collectors. ODIS-RPW is also a probability based destinating mail sampling system used to collect volume information where data collectors record mail characteristics from sampled mail pieces. Since the approval of Proposal Three in Docket No. RM2015-11 in Commission Order No. 2739 (September 30, 2015), ODIS-RPW data collectors enter mail characteristics from digitally captured images of letter- and card- shaped mail from Delivery Barcode Sequence (DBCS) second pass operations, eliminating the need for manual sampling of DPS letters and cards.

B. Use of RCCS Digital DPS Data in Cost Attribution

Total accrued costs for rural carriers are summarized in Cost Segment 10 (CS 10). The costs are divided into separate components for evaluated routes and other routes, based on payroll records.

The route factors are measured during the National Rural Mail Count, which is usually conducted annually. During the National Rural Mail Count, all mail for a large proportion of the rural routes is counted, and time measurements for other factors are evaluated. Therefore, factors related to volume (volume variable cost drivers) and factors independent of volume (fixed cost drivers) are measured during the National Rural Mail Count.

The volume variable costs of rural carrier work hours are determined by a variability analysis developed in accordance with the evaluated time and factors of workload derived from the rural routes participating in the National Rural Mail Count. Volume variable costs are determined for each of the evaluated and other route components of Cost Segment 10.

Data from RCCS are used to distribute the volume variable costs to classes, products – including Extra Services, and rural evaluation categories. The delivery portion of RCCS (data collected via the CODES data collection system) provides mail category data for distribution of the volume variable mail delivery costs.

C. Statistical Study Design

The universe under study in RCCS Digital DPS is all DPS mail being delivered on rural letter routes in ZIP Codes that exist in the ODIS-RPW Digital frame. For rural routes in ZIP Codes not in the ODIS-RPW Digital frame, RCCS manual sampling of DPS mail at the route level will still continue. A two-stage sample design is used for RCCS Digital DPS. The details for each of the stages are described below.

Starting in PQ1 FY18, the RCCS sampling frame was changed to the Universal Delivery Statistics File (UDSF) from the previous Rural Master Pay File. This change was made in order to align the RCCS-Digital frame with the ODIS-RPW Digital frame. While the Rural Master Pay File lists all rural routes, in some cases the ZIP Code did not match that of the Mail Exit Point (MEP) used by ODIS-RPW. The UDSF, having been developed from the Address Management System (AMS) lists all rural routes, and ZIP Codes do align with the ODIS-RPW frame.

First Stage Sample

The ODIS-RPW Digital sample selection process provides a systematic random sample of ZIP-days within each stratum. The ODIS-RPW selection process ensures both geographic and temporal dispersion of the sampled ZIP-days. The use of End-of-Run DPS totals ensures proper weights are used to produce national estimates.

Second Stage Sample (Mailpiece)

The second stage sampling unit is a mailpiece. A systematic digital sample of DPS letter images is obtained, and data from these images are entered by ODIS-RPW data collectors. A subset of these data destined for rural routes only are then processed and expanded to the ZIP-day level using End-of-Run DPS totals for that ZIP.

Relationship to non-digital RCCS

After creating the RCCS sample for each quarter, rural routes are identified that are not included within ZIP Codes covered by ORPW-Digital testing, and therefore not estimated by RCCS-Digital. The DPS mail volumes for these routes will continue to be estimated by the existing manual sampling procedures in non-digital RCCS tests. Volume estimates from both systems are added together to obtain the final RCCS estimates.

D. Estimation and Variance

RCCS produces two types of estimates - volumes and distribution factors (distribution key ratios). A description of the estimates is provided in the overview. The estimates are computed on a quarterly and annual basis. The annual volume estimates are the sum of the quarterly estimates. This section provides the formula for the weighting factor. The formulas for variance, covariance and distribution factors are the same as for non-digital RCCS. They are located in the USPS-FY17- 23_RCCS_Preface_Final document starting on the fifth page.

Notation:

h postal quarter

s skip utilized on record (first stage weight)

T total Rural EOR volume for all ZIP Codes in the digital frame for postal quarter

d weighting factor (second stage weight)

t sampled total volume for postal quarter

n number of sampled pieces on test

v Rural EOR mail volume for the tested ZIP Code on the test date

i compensation category domain

j product or rate domain category

z ZIP-day

l route-day (used for manual sampling weighting)

w weighting factor (used for manual sampling weights)

The weight applied to each record consists of two parts. The first stage weight, indicated by s, is the implied skip. This is calculated by dividing the total Rural End of Run volume in a tested ZIP by the number of RCCS pieces sampled in that ZIP. The second stage weight is indicated by T_h/t_h. This is the total volume for the postal quarter in the RCCS digital frame divided by the sampled total volume.

$$s = \frac{v}{n}$$

$$t_h = \sum v$$

$$d_h = \frac{\text{s x } T_h}{t_h \text{ x } 1000}$$

Variate is defined as follows:

$$y'_{hijz} = \begin{cases} y_{hijz} \text{ if the unit is in the } i^{th} \text{and } j^{th} \text{ domain} \\ 0 \end{cases}$$

The quarterly volume for the digitally collected mailpieces for the j[th] product is

$$\hat{Y}_{hij} = \sum_z d_h y'_{hijz}$$

The total quarterly volume for the j[th] product is the sum of the quarterly volume of the digitally collected mailpieces and the quarterly volume of the manually sampled mailpieces.

$$\hat{Y}_{hij} = \sum_z d_h y'_{hijz} + \sum_k \sum_l w_{hk} y'_{hijkl}$$

E. Processing ODIS-RPW Digital Data

For each ODIS-RPW Digital test, an Image Attribute File is created that provides details that include the destinating carrier route. The Image Attribute files are downloaded bi-monthly and concatenated as a quarterly file to be merged with ODIS-RPW digital data at a later step.

Image Attribute Read In (PC SAS)

INPUTS

Bi-Monthly Image Attribute Files pulled from SPView

OUTPUTS

Quarterly Image Attribute File - ALDRAN.DIG.IAR.FYyy.PQq

ALDRAN.FYyyyy.CHECKIN.JOBS(CHCKNDIG) Weekly mainframe programs copy ODIS-RPW digital data and save to a format consistent with RCCS processing.

INPUTS

HSISMN.ORPW.WEEKLY.DATA.FYyymmwk

OUTPUTS

ALDRAN.DIG.WEEKLY.DIGyymmwk

ALDRAN.FYyyyyQq.RURALDIG.CNTL(ALD501) concatenates the 5 weekly files to create a monthly file.

INPUTS

ALDRAN.DIG.WEEKLY.DIGyymmwk

OUTPUTS

ALDRAN.DIG.MONTHLY.DIGyymm

Control Totals

At the end of each PQ, control totals are calculated based on the EOR data from Network Operations Data Mart (NODM). All rural mail volumes for ZIP Codes in the RCCS digital frame are totaled for each day in the quarter. These totals are used in the expansion of data in a later stage.

Quarterly Volume Estimates and Distribution Factors

Once the rural carrier Digital DPS data for an entire quarter have been validated, quarterly volume estimates and distribution factors are produced. The estimated volumes are compared with the same quarter from the previous year and with estimates from other statistical systems. Substantial differences between the reports are investigated for additional quality assurance.

Quarterly volume estimation is a four-step process. First, monthly files are concatenated to form the quarterly file. Second, the weights used in the estimation procedures are produced. Third, delivery volume estimates are calculated. Fourth, the Z file is produced. The quarterly estimation programs are as follows:

ALDRAN.FYyyyyQq.RURALDIG.CNTL(ALD599) is run to concatenate monthly files to form the quarterly file, merge in the quarterly Image Attribute File, filter out only records destined for rural carrier routes, and assign correct delivery dates and RCCS bucket number to each record.

INPUTS:

Image Attribute Report – DSN = ALDRAN.DIG.IAR.FYyy.PQq
Validated Monthly Data Files – DSN=ALDRAN.DIG.MONTHLY.DIGyymm

Example for FY18 month 10: ALDRAN.DIG.MONTHLY.DIG1810

Only those tests that actually belong in the quarter (indicated by the first digit of the testid) are used for estimation. Below is a list of the months that should be used as inputs for the estimation for each quarter:

PQ1 includes months 10, 11, and 12.
PQ2 includes months 01, 02, and 03.
PQ3 includes months 04, 05, and 06.
PQ4 includes months 07, 08, and 09.

OUTPUTS:

The SAS dataset
DSN = ALDRAN.DIG.RURAL.SHAPE.FILE.FYyyyyQq
with SAS members SHAPE and TESTCNT. SHAPE contains all of the data records for the quarter and TESTCNT includes a listing of all test identification numbers for the quarter.

ALDRAN.FYyyyyQq.RURALDIG.CNTL(DKEYA1) produces first-stage and second weights to be applied to the data received from the ALD599 program.

INPUTS:

Quarterly Shape File - ALDRAN.DIG.RURAL.SHAPE.FILE.FYyyyyQq

RCCS Frame - ALDRAN.PS401T01.RURAL.PQqFYyy

RPW Digital Letter Frame -
HSISMN.ORPW.DIGITAL.FRAMDATA.FYyyQTq

NODM PQ Extract Files - ALDRAN.DIG.RURAL.EOR.PQqFYyy

OUTPUTS

Flat file with weights for processing data

DSN=ALDRAN.RURALDIG.WEIGHTS(FYyyyyQq)

SAS file with weights for processing data

DSN=ALDRAN.DIG.RURAL.FYyyyyQq.YTDWGT.DATA

ALDRAN.FYyyyyQq.RURALDIG.CNTL(DKEYB1) processes delivered mail counts. It merges the mail category information from the mailcode file onto the weighted mail counts file from DKEYA1. This programs also produces the z-file for digital data for the postal quarter.

INPUTS

File with weights

DSN=ALDRAN.DIG.RURAL.FYyyyyQq.YTDWGT.DATA

File with mail category information for the mailcode output file

DSN=ALDRAN.FYyyyyQq.SORTED.MAILCODE(RURLDIG1)

OUTPUTS

Weighted data for each mailcode

DSN=ALDRAN.FYyy.RURALDIG.Qq.MCODE

Weighted data for each CRA bucket

DSN=ALDRAN.FYyy.RURALDIG.Qq.CRABKT

Quarterly Z File

DSN=ALDRAN.RURALDIG.Z.FYyyyyQq

After RCCS Digital and Regular RCCS have completed processing of the respective MCODE and CRABKT files, they are read into a PCSAS program that combines the Digital and non-Digital DPS estimates to produce the final volumes and distribution factors in the quarterly RCCS Matrix spreadsheet.

Annual Estimates

Annual volume estimates are used to distribute costs to categories of mail. The volumes are calculated by summing the quarterly volumes. The annual volumes program is executed from the following file: ALDRAN.FYyyyy.RURALDIG.CNTL.

ALDRAN.FYyyyy.RURALDIG.CNTL (RDIGMCOD) is used to produce annual rural digital delivery mail volume estimates by mailcode.

INPUTS:

The quarterly volumes files:

DSN=ALDRAN.FYyy.RURALDIG.Q1.MCODE

DSN=ALDRAN. FYyy. RURALDIG.Q2.MCODE

DSN=ALDRAN. FYyy. RURALDIG.Q3.MCODE

DSN=ALDRAN. FYyy. RURALDIG.Q4.MCODE

OUTPUT:

The annual volume file for digital DPS delivered mail volume estimates by mailcode.

DSN=ALDRAN. RURALDIG.FYyyyy.MCODE.DATA